SEEDS
Sustainability Environmental Economics and Dynamics Studies

Working Paper Series

Regional Innovation Systems in China: A long-term perspective based on patent data at a prefectural level

by

Giorgio Prodi, Federico Frattini, Francesco Nicolli

03/2016

SEEDS Working Paper 01/2016
April 2016
by Giorgio Prodi, Federico Frattini, Francesco Nicolli

Regional Innovation Systems in China: A long-term perspective based on patent data at a prefectural level

Giorgio Prodi[a] Federico Frattini[a*] Francesco Nicolli[ab]

[a] Department of Economics and Management, University of Ferrara, Ferrara, Italy; [b] Research Institute on Sustainable Economic Growth, National Research Council, Milan, Italy; [*] Corresponding author. Department of Economics and Management, University of Ferrara, Via Voltapaletto 11, 44121 Ferrara FE Italy. Email: federico.frattini@unife.it

Abstract

This paper focuses on the connections between long-term development and Regional Innovation Systems (RIS) in China. It aims to investigate how the evolution of RIS fits with China's overall process of economic upgrading. The analysis relies on Chinese patent applications filed to the EPO during the period 1981 to 2009, which authors have regionalised at a prefectural level. Conceptually, the investigation concerns the relative prevalence of indexes derived from inventors' and applicants' localisation to describe local innovation activities in terms of emergence, development and reinforcement. The hypothesis ranks higher those prefectures where indigenous applicants prevail, that is, the initiative, organisation and exploitation of innovation activities are foremost local (or endogenous). Results return the possibility of grouping Chinese prefectures into six clusters. On this basis, RIS features appear to diffuse, even while regional concentration of innovation activities is still increasing. This pattern is deemed to fit the process of industrial development in China very well. As it was in the past, RIS benefit from the opportunities that a long-term development strategy provides, but face its limits as well.

Keywords: China; development; endogeneity; patent; reform; Regional Innovation System

1. Introduction

Regional Innovation Systems (RIS) are defined as places where innovation activities are concentrated and networked, creating a context-specific environment that fosters the production of knowledge with a systemic configuration (Cooke, Gomez Uranga, and Etxebarria 1997; Cooke 2001). This topic has recently achieved resounding success with researchers in the field of economic development (Fagerberg and Srholec 2008; Lundvall et al. 2009; Srholec 2011). In fact, innovation systems represent a tool for economic catching-up in middle-income countries (Fu et al., 2011; Lee, 2013), including China (see Fan 2014 for a survey). However, an inquiring perspective that narrowly focuses on innovation activities and policies is poorly promising for emerging economies, because the rise of RIS in this context is structurally coupled to the enhancement of broader capabilities and a process of institutional change. Accordingly, this paper investigates the evolution of regions towards RIS and its connection with the long-term development process in China.

This linkage is particularly crucial when technological upgrade is the outcome of strongly unbalanced processes that need to be turned into widespread development (Gu and Lundvall 2006). Compared to other middle-income countries, economic growth has been very impressive in China during the last three decades suggesting that a distinctive approach was driving the change. Such a growth cannot result from capital accumulation only but, more plausibly, also a mix of structural change, economic transition and industrial development (Naughton 2007). Nonetheless, there are well known gaps across Chinese regions in terms of economic restructuring. Hence, focusing on RIS and their structural features can provide valuable insights on the accomplishments in regions that represent the country's development frontier, as well as the opportunities of catching-up for latecomers. For this reason, regional innovation paths are discussed here considering a broader development framework.

Section 2 briefly recalls the main development steps since 1978 in China, highlighting the key elements in the long-term development strategy at a national level, pointing three main reform stages and stressing the role of economic restructuring and transition in creating the conditions for growth. Are these conditions also suitable for a more recent step represented by RIS? They involve fundamental capabilities at a local level. Are regions properly endowed? Unlike the provincial level that is frequently observed in literature, this paper considers the prefectural one to better understand the heterogeneity that the Chinese development process entails. Unfortunately, at this regional detail, the lack of data prevents a comprehensive examination of the economics in the long term. Therefore, Section 3 introduces an empirical approach that points directly at capturing the local endowment of capabilities, based on a clustering procedure. Variables entering the analysis are innovation-related indicators arranged at a prefectural level from the Chinese patent applications to the European Patent Office (EPO) collected in the OECD REGPAT database (January 2014). More precisely, the information about the inventors and applicants' location are mixed to draw near the embeddedness of regional innovation activities and understand its consistency with the national development strategy. Then, Section 4 reports the results from a cluster analysis and compares them to some main fact in the geography of development in China, especially in terms of stage-by-stage evolution. At final, Section 5 provides a few considerations about open issues.

2. Historical background and research question

Economic growth in China is often considered as a main consequence of an increase in exportations. However, literature showed that export was only one among many driving forces (D. He and Zhang 2010) that are based on improvements in factor productivity and technological sophistication (Rodrik 2006; S. Chen, Jefferson, and Zhang 2011). Hence, China's growth has been primarily fostered by a general enhancement of industrial capabilities in a progressively closer interaction with international markets (Brandt, Rawski, and Sutton 2008; Felipe et al. 2013). This represent a pillar since 1978, when China committed to a gradual transition to the market economy. The leading force in this process was not an invisible hand but was rather a governmental action reforming rules and institutions to support structural change. The nexus between the institutional and economic side is a set of industrial policies designed to drive the country during a slow and experimental integration into international markets. Along this development path, China undoubtedly demonstrated the ability of catching up in industrialization. However, stepping forward to the knowledge-based economy is not a consequential task. RIS actually represent an important piece in this change (X. Li 2009) and a main argument in redesigning the governmental intervention to attract foreign technologies (C. Chen, Chang, and Zhang 1995), establish science parks (Hu 2007) and support the creation of spin-offs (Kroll and Liefner 2008).

2.1. The Chinese development strategy after 1978

Post-Mao reforms in China can be split in three periods setting as much stages in the transition from a planned to a market economy (Naughton 2007; Brandt, Rawski, and Sutton 2008). The first stage started in 1978, contextually to the political debate about a new pace of social and economic development. The opening step was an early dismantling of government control over the economy and the creation of fundamental market players, i.e., private firms. The most important initiatives were: defining a dual track regime that allowed State-Owned Enterprises (SOE) to partially develop their own businesses outside of a plan; permitting Town- and Village-owned Enterprises (TVE) to run businesses completely outside of a plan; the opportunity for private Small and Medium Enterprises (SME) to produce; allowing foreign firms to operate in China, initially in joint venture with domestic firms but later independently; creating Special Economic Zones (SEZ) as preferential channels for international trade (Naughton 2007; Frattini and Prodi 2013a). Hence, China was no longer hermetically sealed, and the national government implemented consequent actions to normalise the economic relations with foreign countries. Of course, transition was not completed yet (Rawski 1994) and market dynamics were still biased by the residual governmental control over industries and firms (Zhang and Tan 2007; Brandt, Tombe, and Zhu 2013), but the fundamental pillars for Chinese development were established during these years. Very soon, SEZ became places where testing experimental policies and market institutions (Heilmann 2008) and attracting foreign capital and technologies (Fu 2008). Thus, SEZ represented an essential seed for the creation of a Chinese path to technological upgrade and innovation.

During the first stage, the Chinese economy went restructuring, but reforms were cautious regardless and took a 'no loser' approach (L. J. Lau, Qian, and Roland 2000). Conversely, at the beginning of 1990s, the government changed the reform style, and attributed exclusively an addressing role to five-year planning focusing its action primarily on SOEs (W. Li and Putterman 2008). During this second stage, market mechanisms became more robust and consistent with an effective selection process. SOEs are currently bigger, more capital- and knowledge-intensive, more productive and able to profit (Gabriele 2010). Moreover, the increase in Foreign Direct Investment (FDI) strongly supported industrial development, and it stimulated upstream and downstream productive connections (S. Sun 2012), providing the domestic industry with new competitive pressures (Brandt and Thun 2010) and new technological concerns (Girma, Gong, and Gorg 2008). Combined with the accumulation of physical and human capital and the protection of infant industries, the activities of foreign Multi-National Enterprises (MNE) pushed the diffusion of innovation activities at a regional level, encouraging a general upgrading of manufacturing technologies as well as the access to international supply chains (Y.-C. Chen 2007).

At final, China's World Trade Organization (WTO) membership in 2001 opened a third stage of reforms, which reinforced the transition and internal market growth. Incentives to attract foreign firms were reduced and shifted to empowering the indigenous contributions to economic development. Some industries were identified as being strategic to promote the growth of national champions (Hemphill and White 2013) and acquire relevant assets in foreign countries (Deng 2009). Nonetheless, on the formal side, these steps forward in economic transition do not represent an ultimate choice of capitalism because market is relevant in China, but merely as a way for development. Necessarily, this vulnerable equilibrium has further pressured institutions to change, especially in terms of law enforcement and protection of intellectual property rights (Wu 2007b). Thus, China's catching-up process is experiencing a new phase leading the country toward the knowledge-based economy (Dahlman and Aubert 2001).

2.2. Exogenous seeds, tumultuous growth and innovation

According to the progression of reforms, the 'preferential policies' in the Coastal area and a simultaneous dramatic increase in patenting (Figure 1), an early significant step in diffusion of technological activities in China occurred at the end of the second stage. Actually, patents are usually adopted as a proxy for the output of innovation activities (Keller 2004). At that time, FDI and MNE were favoured tools for promoting industrial development, and SEZ were an important doorway to access innovation and technology in developing economies (Lall 1992). However, literature often uses a very broad notion of SEZ that inaptly includes different policy initiatives (Zeng 2010). In China's case, SEZ in the strict sense were established since early 1980s in Shenzhen, Zhuhai, Shantou, Xiamen and Hainan, and later, they spread to many other areas (Shanghai Pudong New Area and Tianjin Binhai New Area are the most renowned among them). More precisely, SEZ are geographically delimited areas 'with a single management or administration and a separate customs area (often duty free), where streamlined business procedures are applied and where firms physically located within the zone are eligible for certain benefits' (World Bank 2010). In addition, in these areas, the development of technology- and innovation-related activities is frequently supported by complementary policy initiatives, such as industrial, technological and science parks, which are defined as agglomerations of physical infrastructures in the high-technology domain. This physical component is combined with functional ones, such as specific knowledge, services and financial providers, creating new business opportunities and adding value to mature companies, fostering entrepreneurship, incubating new innovative companies, generating knowledge-based jobs, and building attractive spaces for knowledge workers' (World Bank 2010).

[Figure 1 about here]

Unlike spontaneous clusters and other unstructured agglomerations of firms, SEZ can be considered as exogenous seeds of industrial development. They primarily aim at attracting external investments, technologies and knowledge during the early steps of upgrading, while a clustering process of local activities is emphasised later only (Frattini and Prodi 2013b), especially in least protected industries (C. He, Wei, and Xie 2008). This policy approach was downsized during the third stage, given the shift from a 'defying' to a 'following' comparative advantages approach in the national development strategy (Lin and Wang 2012). During the same period, the number of patent applications from China at the EPO literally exploded, prompting a technological emancipation in some regions.

Chinese clusters today are mainly localised in the Eastern area, but agglomerations are also moving towards the Southeast. The largest are mainly located in the provinces of Jiangsu, Shanghai, Zhejiang, Fujian and Guangdong, areas that began growing first. In these regions, the Gross Domestic Product (GDP) per capita is generally higher than the national average, and industrial capabilities already include technological and innovation activities. Agglomeration processes matched the diffusion of innovation activities at a regional level in terms of both intensity within and distribution across regions (Figure 2). Thus, a general upgrading of industrial capabilities has progressively involved a wider range of areas, but it did not prevent the most capable regions from continuing to increase the frequency of their patent filing to the EPO, i.e., the volume of innovation activities. Evidently, this is another side of the regional disparities pushed by a very tumultuous growth.

[Figure 2 about here]

2.3. Research question

As mentioned, RIS are regions where innovation activities are systemic and supported by well-structured processes for governing, financing and learning (Cooke, Gomez Uranga, and Etxebarria 1997). Hence, RIS are multi-layered arrangements of spatially concentrated activities (Srholec 2011) based on the connections between firm-, government- and research-type organizations (Cooke 2001) that generate contextual effects on the production of knowledge and innovation (Kaihua Chen and Guan 2012). According to this definition, this paper focuses on the progression of innovation activities across Chinese regions in terms of embeddedness and emancipation from external initiatives and funding. The absorption of imported technologies is an essential step to start innovation capabilities at a local level (A. K. W. Lau and Lo 2015). However, it could slow down industrial upgrading if complementary indigenous aptitudes are not developed (Yifei Sun 2002; Fu, Pietrobelli, and Soete 2011).

Attracting investments and technologies to the Coastal area was for long time a priority in the Chinese strategy to foster industrial upgrading. Consequently, many relevant seeds of innovation activities were primarily exogenous, except for a few cases such as Beijing, where important research and academic institutions were already functioning as catalysts of external knowledge (Kun Chen and

Kenney 2007). For this reason, the paper attempts to understand whether the steps in industrial upgrading promoted by the government strategy also turned into steps towards RIS at a regional level. In other word, the aim is to understand if and how this strategy was able to set the conditions for the rise of RIS within a more general process of capability enhancement, highlighting the connections between patenting diffusion and development timing.

To grasp this point, let consider localisation of patents. A way to measure the regional contribution to patenting is to identify the inventors' location: inventors are the human capital embodying the essential innovation capability. Nonetheless, inventors do not provide information about funding efforts or the exploitation of their inventions, which are an essential part of RIS configurations. Hence, observing the location of applicants, who are generally appropriating the innovation returns, can better capture the location of innovation initiatives. Combined, these two perspectives allow a more detailed picture about the distribution of innovation and technological capabilities at a regional level. The differences between the information provided by inventors and applicants are generally larger in those regions where the intensity of innovation activities is poor. However, such differences in China are rather relevant also observing the most patenting prefectures (Figures 3(a) and 3(b)). In particular, the prominence of foreign applicants, which here embody external initiatives, has recently decreased, but is still considerable in 2009. Conceptually, the main assumption here is RIS being regions where the contribution to and the exploitation of inventions are reciprocally balanced or where the exploitation prevails, so that they are active players in promoting innovation and technologies. In this sense, the sources of technological capabilities are mainly endogenous.

[Figures 3(a) and 3(b) about here]

Generally, this case is the result of spontaneous clustering phenomena and, especially in China, policy actions. Nonetheless, both depend on intrinsic development processes at a local level, including be endowed with economic resources. For instance, many policy initiatives have affected the diffusion of patenting activities in China. In particular, provinces and cities launched patent subsidy programs since 1999, which evidently contributed to the surge of patents applications to the Chinese patent office and abroad (X. Li 2012). Clearly, to be successful, these strategies needed a proper resource endowments to subsidise and indigenous capabilities to be subsidised.

Starting from those premises, the paper focuses on different properties of patenting deduced from the location of inventors and applicants. The early diffusion of innovation activities across regions is expected to particularly involve aspects related to the inventors. Conversely, a higher level of embeddedness and endogeneity should presumably emerge during the third stage, when the importation of foreign technologies becomes relatively weaker and spillovers more relevant on both the technological and economic sides.

3. Research setting

3.1. Patent data at a prefectural level

Technology is a complex phenomenon and a comprehensive description is very difficult regardless of the measure adopted. One possible empirical solution is to look at the output of innovation activities and related capabilities, i.e., patents (Keller 2004). Thus, as introduced, the analysis relies on patent applications from China to the EPO over the period 1981-2009, which are publicly available in the OECD REGPAT database (January 2014).

Although patents have several shortcomings as an indicator of technological capabilities, they also have several advantages. First, there are very few examples of economically significant inventions that have not been patented (Dernis and Guellec 2001). Second and more importantly, a strong correlation to R&D spending makes patent data a good proxy for innovative activities (Griliches 1990). Third, patent statistics could perform even better than technology input measures for evaluating the innovation capabilities across regions in the case of China (Guan and Liu 2005). Fourth, focusing on patents filed to the EPO among all of the different types of patent documents has its own advantages. One concerns the expectation that, based on the higher costs of seeking protection abroad compared to filing documents solely at the home countries' patent offices, patents at the EPO would generally reflect high-quality inventions with a more homogeneous economic value. Moreover, a technology that is 'new to the country' is not necessarily 'new to the world' (X. Li 2009). Thus, referring to the EPO rather than to a national patent office works primarily as a quality threshold (Dernis and Khan 2004), also when studying innovation activity outside of the European Union (EU) as in this paper.

A conceptual concern may however arise about the appropriateness of using EPO instead of SIPO (State Intellectual Property Office of the P.R.C.) data in the case of China. This issue could be relevant in part before 2002, when the number of Chinese patent applications to the EPO was very low. Nevertheless, it is important to stress that, according to the perspective here adopted, a low patent count to the EPO is a result *per se*. It highlights a low interest in seeking protection abroad likely due to an unsuitable value of inventions in terms of both technology and profit, such as they could not generate appropriate returns to compensate for the opportunity cost of filing a patent outside of national borders. In other words, moving from null to positive values of EPO applications from China is not just a matter of commercial strategy but rather it concerns a change reflecting innovation activities growing more competitive at the global level, which is itself an indicator of a RIS configuring process.

Even though the theoretical consistency, on the empirical side, the robustness of using EPO data can be tested using the correlation between the applications from China to the EPO (catalogued according to the 'China' country attribution) and SIPO (catalogued as 'domestic'). This simple empirical exercise is carried out in Appendix A showing that all correlations tend to be very high and supporting the choice of relying on EPO documents only. Moreover, tests of correlation sustain the fair convenience of excluding specific regions from the data set. Taiwan, Hong Kong and Macau, in fact, followed their own history, autonomously with respect to China. Accordingly, these regions are potential sources of bias in the analysis and, therefore, dropped from the data set.

In addition, the analysis of innovation-clustering phenomena can benefit from a regional classification as much detailed as possible, given that these processes are enhanced by systemic interactions among innovators (Srholec 2011) and region-specific factors are particularly relevant in China (Wang and Lin 2012). For this reason, the provincial level the OECD database includes for China is very coarse, and data have been rearranged with a new prefectural-level attribution based on a semantic search in the 'address' field associated with each inventor and applicant from China (Callaert et al. 2011). Once controlled for the initial provincial-level attribution in the OECD database, the outcome is a data set of 20,202 patent applications from 1981 to 2009 distributed over 200 of the 345 prefectures considered, which includes all patents with at least one inventor or applicant from China. As an extension, this procedure allowed adjusting some mismatches in the OECD database, that affect 2.4% of all inventors and 2.1% of all applicants considered. Then, the information related to each single patent document have been aggregated by inventor' region, applicant' region and priority year through a fractional count (OECD 2009). As a result, the analysis presented here makes use of a full-country panel of data over 30 years that includes several null values, but is novel and more detailed than usual in literature.

3.2. Methodology

A cluster analysis is an effective method for disentangling different patterns of innovation activity across Chinese prefectures, especially using reforms periodisation to simplify the temporal dimension of data. Cluster analysis aims ‘to classify a sample of entities (individual or objects) into a small number of mutually exclusive groups based on their similarities’ (Hair et al. 2009, 20). In this paper, the empirical procedure focuses on similarities in the distribution of three indexes created to describe qualitative features in the clustering process of innovation activities at a local level: the prevalence of patents with indigenous inventors only (INV_{is}), indigenous applicants only (APP_{is}) or both indigenous inventors and applicants ($BOTH_{is}$) in each prefecture during each reform stage.

Indexes have a common denominator (tot_{it}) defined as the sum of all the patent fractions attributable to the prefecture i in year t without distinguishing between applicants and inventors. Then, the number of patents identified by applicants (app_{it}), inventors (inv_{it}) and their overlap ($both_{it}$) enter the respective indexes as the numerator. More precisely, indexes, identified by capital letters, are calculated as follows:

$$APP_{it} = (app_{it} - both_{it}) / tot_{it} \quad (1)$$

$$\mathrm{BOTH_{it}} = \mathrm{both_{it}} / \mathrm{tot_{it}} \quad (2)$$

$$INV_{it} = (inv_{it} - both_{it}) / tot_{it} \quad (3)$$

where $tot_{it} = app_{it} + inv_{it} - both_{it}$, so that for each prefecture i where $tot_{it} > 0$ in year t

$$APP_{it} + BOTH_{it} + INV_{it} = 1 \quad (4)$$

However, all the indexes assume an indefinite form when $tot_{it} = 0$ in a given here t. In this case, to do not jeopardise the observations collected (200 prefectures), a null value is assigned to each index so that it can enter the clustering procedure. Conversely, when $tot_{it} = 0$ for all the years t within a reform stage s, the observation i is excluded *ex ante* to let the analysis better discriminate across groups of prefectures with a positive patent count in the period, while prefectures formerly ruled out can be retrieved *ex post* in an arbitrary cluster to complete the country map (345 prefectures).

The dispersion of indexes values across prefectures suggests that they consistently vary in time and space on a yearly base. Accordingly, they are supposed to capture different dynamics at a regional level (Figures 4(a), 4(b) and 4(c)). Nonetheless, yearly changes can be very relevant, especially when patent counts are very poor. For this reason, only within-stage averages *s* for each prefecture *i* enter the cluster analysis. Naming X_{it} the yearly value of whatever index, the related period average X_{is} is calculated as:

$$X_{is}=\frac{1}{T}\sum_{t=1}^{T} X_{it} \qquad (5)$$

where *T* is the overall number of years *t* in a stage *s*.

[Figures 4(a), 4(b) and 4(c) about here]

This simplification provides the clustering procedure with several empirical advantages and a main con. First, when $tot_{it} > 0$, yearly index values represent shares on the total patent applications. Of course, they are unrelated to the patenting intensity and prevent to consider dissimilarities due to the prefectural amount of patents. Differently, the period average allows somehow rescaling indexes based, in this case, on the duration of innovation activities in each prefectures, so that values entering the analysis are higher when prefectures exhibit a longer patenting history. This is even true when patenting traditions are discontinuous, since all three indexes assume a null value for those prefectures, at least once (*t*) within the period *s*. Second, within-stage averages attenuate the randomness of values potentially due to low patent counts entering the indexes as denominators, an issue that is particularly relevant until mid-1990s. However, despite mitigation, the country-mean value of indexes is notably different across periods and substantially increasing in time (Figure 5). Therefore, the solution adopted is expected to improve the procedure adopted to distinguish among different levels in the diffusion and embedding of technological capabilities across prefectures and reform stages. Third, detaching properties of technological capabilities from patenting propensities is very difficult using patents as proxies for technologies, but the indexes presented here appear theoretically robust to this issue. In fact, whether the measurement of patenting propensities is generally based on positive counts, such as the number of patents per inhabitants, total prefectural counts here lose statistical importance as the number of patents increases. Last, despite several advantages, this approach has the main flaw of preventing the analysis from being performed on a fixed set of observations across the three reform stages that includes 40, 87 and 187 prefectures, respectively. For this reason, as a robustness check, Appendix B presents and discusses the same analysis but considering an unvaried number of prefectures across stages. It is obtained letting all of the prefectures that record at least one patent application to the EPO over the entire period (1981-

2009), not a single stage, enter the clustering procedure, and dealing with a noteworthy number of missing values converted into zeros.

[Figure 5 about here]

Conceptually, technological capabilities are expected to increase when the opportunities for committing, funding, managing and exploiting the returns of research activities also increase. In this sense, the analysis attributes to the resulting clusters not a distributional meaning only, but also an ordering value related to the endogeneity of capabilities. The embeddedness of innovation activities is considered to increase, moving from the prevalence of *INV* to the prevalence of *APP* and, consequently, strengthening the conditions for RIS-building processes. Indeed, when *INV* shows the highest value in a group of prefectures, it means that their innovation activities are mainly supported by external initiatives due to a lack of indigenous applicants. Conversely, higher *APP* values suggest that prefectures are able to attract and exploit creative resources located elsewhere, capturing the return of their inventions. Final, *BOTH* is related to activities that are primarily developed at a local level. Although knowledge flows being open is relevant for innovation systems (Bell and Albu 1999), *BOTH* is probably the most appropriate index capturing the embeddedness level necessary to consolidate those capabilities needed to absorb external knowledge and upgrade technological capabilities (Asheim and Vang 2011), while *APP* can be associated more with the appreciable maturity of the innovation environment. Thus, a combination of *BOTH* and *APP* already denote an outstanding endowment of technological capabilities.

4. Results and discussion

Summarising the preparatory steps to the analysis, information about the prefectural location of patent inventors and applicants have been used to generate three indexes: *APP*, *BOTH* and *INV*. They measure the relevance of one or both actor types within prefectures that is associated to a different level of endogeneity in local innovative activities. When missing patent data hinder calculating indexes, all zero values are assigned to prefectures. Moreover, these prefectures are excluded *ex ante* from the cluster analysis when values are zero for all the years in a reform stage. In all of the other cases, yearly values are averaged within each reform stage. Now, to classify the index values finally obtained in a few meaningful groups, the analysis performs a two-step clustering procedure, with a first hierarchical step whose results are used as initial seeds in a second non-hierarchical approach

(Hair et al. 2009; Rizzo, Nicolli, and Ramaciotti 2013). This standard method is applied separately to each one of the three periods.

From a technical perspective, the hierarchical analysis employs the squared Euclidean distance as measure and the average distance as grouping method, while the criterion for the selection of a specific cluster solution is the presence of a significant leap between the values of agglomeration coefficients (Manly 2004). On this basis, the process suggests both seven- and eight-cluster solutions for the first stage, but the seven-cluster solution appears easier to be interpreted with the agglomeration coefficient increasing from 2.3 to 5.4 (Table 1(a)). Conversely, the procedure suggests a clearly identifiable six-cluster solution for the second stage and a five-cluster solution for the third, in which agglomeration coefficients jump, respectively, from 4.9 to 6.4 (Table 1(b)) and from 5.5 to 7.8 after a smooth increase (Table 1(c)). Then, cluster centroids just obtained are introduced as initial seeds in a K-means non-hierarchical cluster procedure that reassigns observations by following iterations until the best performing separation is configured (Hair et al. 2009). This two-step process allows optimising the analysis, given the higher flexibility of the K-means non-hierarchical compared to standard hierarchical technique and, contextually, the solution to the dilemma of choosing proper initial seeds for a non-hierarchical methods that the prior hierarchical tool offers.

[Tables 1(a), 1(b) and 1(c) about here]

Other than a wide diffusion of innovation activities across prefectures, results immediately exhibit three tendencies. First, centroid values generally increase stage-by-stage. It means that, even though the surge of patent applications is primarily due to a few regions, the average intensity of patenting in Europe is generally growing over time in lower-performing prefectures as well. In this sense, results recall the long-term dynamic formerly depicted in Figure 2, but now they sketch a clearer transitional process across stages. Second, the number of groups decreases over time, although it is evident that some clusters include one or two cases only in the first period. Nevertheless, the reduction in the number of groups can relate to an overall reinforcement of innovation activities. In fact, the high concentration of innovation activities during the first reform stage is primarily due to a low number of patent applications to the EPO that induces some observations to stay isolated, representing overly specific cases and not contributing to disentangle a fuzzy picture. Conversely, more intense innovation activities during the third reform stage allow data to more consistently describe effective regional disparities. Hence, results basically suggest that during the three reform stages, innovation activities have spread across prefectures in China, and their intensity has grown over time.

However, comparing the distribution of observations across a set of clusters varying across periods is very difficult and prevents a deeper discussion. For this reason, clusters obtained for each reform stage through the analytical procedure have been converted into a unique descriptive framework structured into six groups named 'adjusted clusters' (*ACL*), also including prefectures that did not enter the analysis because they did not record any patent applications (Table 2).

[Table 2 about here]

Groups in the first and second stage are merged according to the prevalence of indexes, while almost all of the centroids exhibit a satisfying distance in the third stage. This induces to finally fix results in four groups. Then, two more groups are added to complete the description of all possible records: one including all of the prefectures with no patent applications to the EPO and one empty group stressing the lack of regions exhibiting a prevalence of *APP* only. Finally, adjusted clusters have been ordered and numbered from 0 to 5, assuming that the accomplishments in innovation environments progressively moves through a high level of *INV* at the beginning, then *BOTH* and finally *APP*. As previously discussed, the prevalence of *INV* means that innovation activities generally depend on initiatives that are exogenous to the region because who exploits the invention's result is located elsewhere. Instead, high values of *BOTH* imply that applicants and inventors both have the same location, suggesting that innovation activities are well structured in the region and able to generate returns captured *in loci*. Last, the prevalence of *APP* allows supposing that prefectures have gone a step further in also exploiting non-indigenous inventions and, as a consequence, polarising external activities.

No group shows *APP* prevalence until the third reform stage, although combined with *BOTH* only (4), confirming a progression of innovation activities in terms of embeddedness and emancipation. These very few prefectures appear to mix well-structured innovation capabilities and a strong aptitude for attracting and exploiting external knowledge. At the same time, the number of prefectures where *BOTH* prevails (3) increases, including those regions where the conditions for RIS have probably been mostly built but still need to further develop and turn into crucial junctions for external knowledge and technological flows. On the other front, the number of prefectures filing patents to the EPO dramatically increases over the reform stages, growing the group of regions where *INV* prevails (1) from 29 to 130 prefectures. This broad diffusion of exogenously driven activities is partially compensated by those prefectures shifting into groups where capabilities are becoming emancipated, mainly during the third stage, and *INV* and *BOTH* prevail together (2).

Therefore, results reflect two complementary phenomena that geographical maps can help to grasp (Figures 6(a), 6(b) and 6(c)). Innovation activities diffuse across prefectures, mainly exploiting external seeds (1-2), while innovation capabilities appear to be well structured and embedded only in few prefectures (3-4). For these, RIS either already exist or, at least, the conditions exist that will lead to RIS. Consequently, adjusted clusters (*ACL*) provide a coherent picture of a transitional process that turns each step forward into a new combination of development determinants. The 'dualism' returned is just a part of this transitional context, where this and other types of 'duality' couple with both the impressive increase of innovation activities and their high variability across regions (X. Li 2009). Among them, the main 'dualism' in this paper concerns the coexistence of different capability endowments marking the distinction between substantial endogenous and exogenous innovation processes.

[Figures 6(a), 6(b) and 6(c) about here]

Consistent with the development history in China, in the early stage of economic development, the highest levels of innovation capabilities were located in few well-populated urban sites. Although still poor in terms of intensity, innovation activities started growing and diffusing during the second reform stage, especially in the Coastal area, anticipating what will have more widely happened during the last stage. In fact, the increase in number of prefectures filing patents to the EPO after 2001 is impressive. Despite that, the connection between the step towards higher levels of innovation process embeddedness and the location of innovation activities during the previous periods is quite clear. With very few exceptions such as Chongqing, adjusted clusters *ACL* 3 and 4 include prefectures sited in the Coastal area and, accordingly, the map of adjusted clusters quite effectively overlaps the more general development scenario in the country. Actually, the intensity of patenting activity, supported by the degree of development, can play a positive role in strengthening the properties of innovation environments. However, patents are not only a matter of R&D effort but also of efficiency and incentives in the innovation processes (X. Li 2009; X. Li 2012), 'infra-structural' and 'super-structural' components (Cooke 2001), formal and informal interactions and networking (Cooke, Gomez Uranga, and Etxebarria 1997), and the ability to exploit knowledge spillovers (A. K. W. Lau and Lo 2015). These elements are generally the same in the advanced and emerging economies, but the capabilities that each protagonist exhibits are often not comparable and, in the last case, a potential source of weakness (Kaihua Chen and Guan 2011). Nevertheless, the Chinese development process appears to have been able to effectively produce the conditions for RIS in the long term, at least in a few prefectures where the *BOTH* and *APP* indexes prevail. This evidence answers the first part of the research question, but the analysis still misses an essential step to understand how these conditions were created.

The paper mentioned some regions as exemplificative places for exogenous seeds (the SEZ) and for endogenising capabilities (the industrial agglomerations). Table 3 compares the distribution of these regions across adjusted clusters (*ACL*). The SEZ-hosting prefectures have been assigned to different clusters since the first reform stage including those in which *BOTH* prevails (3), although they are mainly concentrated in the groups where *INV* only prevails (1). Interestingly, they show the lowest rate, with no activity at the EPO (0) in the period 1981-1992 with respect to other provinces, but several of them immediately moved toward better innovation environment conditions (2, 3) between the first and second stage. This dynamic is consistent with previous discussion, where SEZ have been defined as 'doorways' for innovation and technological capabilities. However, the seeds that are exploitable in these cases are also primarily considered exogenous, and this is likely the reason why no appreciable enhancement of capabilities emerges between the second and the third reform stages. Conversely, innovation activities in regions where industrial activities agglomerated, such as Fujian, Guangdong, Jiangsu and Zhejiang, are absolutely weaker before 1993 and most of the prefectures enter the group of no activity (0). Moreover, the transition toward higher adjusted clusters between the first and second stage is generally slower here than in the case of SEZ, but in a few prefectures, innovation capabilities were able to further improve in the following step, even joining the groups where both the *APP* and *BOTH* index prevail (4).

[Table 3 about here]

Again, this picture is coherent with the long-term development process summarised in Section 2, although excluding SEZ from provinces in describing the distribution of adjusted clusters undoubtedly emphasise it. Moreover, SEZ essentially contributed to foster innovation activities and generate spillovers that surrounding regions could have exploited. Indeed, main industrial agglomeration areas host SEZ, as in the case of Guangdong (Shantou, Shenzhen and Zhuhai) and Fujian (Xiamen).In other cases like Jiangsu and Zhejiang (Shanghai), they are in close geographical proximity with SEZ they are close to other regions not included in the analysis, such as Hong Kong (Guangdong) and Taiwan (Fujian). Thus, a strategy aiming to open delimited 'doorways' to the market economy during the first reform stage, outlining transition and industrial upgrading, appears to have effectively worked also for innovation activities and technological capabilities.

5. Final considerations

This paper aimed at investigating how the long-term development process in China, which is closely related with a national government strategy of industrial upgrading, provided some regions with the capabilities needed for being RIS. They are regions where innovation capabilities are well-structured and embedded at a local level (Cooke, Gomez Uranga, and Etxebarria 1997). Accordingly, the empirical strategy focused on patent applications to the EPO from Chinese inventors and applicants to derive three indexes aiming to capture how much innovation activities would be or become endogenous in prefectures. The analysis returned different clusters that highlight the progressive diffusion, evolution and embedding of capabilities through the development stages. Further, it delimited main changes in the governmental approach and the main steps in the transition from a planned to a market economy (Naughton 2007; Brandt, Rawski, and Sutton 2008).

Results show that RIS are nested in this wide process, although they substantially emerge after 2001 only and are still located in very few prefectures. The analysis presented does not claim to provide other more specific insights about the innovation activities conducted at a regional level, such as their technological specialization or orientation, institutional nature and, especially, geographical boundaries. In fact, the research question posed in Section 2 can be validated only through a comprehensive view of China in terms of both the spatial and temporal dimension. In this sense, the long-term perspective from 1981 to 2009 and the geographical detail at a prefectural level represent a value added in the analysis. In particular, this approach provides evidence for the crucial role of SEZ in seeding regional innovation activities through the importation of foreign technologies and their diffusion in the surrounding areas. However, at the same time, the analysis provides clear evidence for the later involvement of the internal provinces in those development processes that first drove economic growth in the East and Southeast.

For this reason, in China the RIS scenario appears to fit the stages of economic transition, benefiting from their opportunities but also paying costs in terms of regional gaps. Therefore, a main risk is that the creation of a RIS network in China could nurture a recursive process of polarisation similar to the previous that the government started fighting at the end of 1990s, also in terms of technological disparities, with the so-called 'Go West' strategy (Tian 2004). In fact, results show that patent applications to the EPO have been spreading across the central regions during the last reform stage. Thus, they suggest that national interventions are still essential in governing this tumultuous process. However, the increasing diffusion of innovation activities contextually asks other institutions such as regional authorities, banks and universities to play a leading role (Wu 2007a; Kaihua Chen and Guan 2011; Zhao et al. 2015). They should be protagonists in supporting the embedding process of technological and creative capabilities, particularly making regional innovation environment more centred on private funding (Yutao Sun and Liu 2010). Consequently, a policy approach based on closer interactions between multiple institutional levels is likely to be a useful step toward a more convergent development process but, unfortunately, institutional capabilities are even harder to build than innovation ones.

References

Asheim, Bjørn T., and Jan Vang. 2011. "Regional Innovation Systems in Asian Countries: A New Way of Exploiting the Benefits of Transnational Corporations." *Innovation: Management, Policy & Practice* 8 (Innovation in China: Harmonious Transformation?). eContent Management Pty Ltd: 27–44. doi:10.5172/impp.2006.8.1-2.27.

Bell, Martin, and Michael Albu. 1999. "Knowledge Systems and Technological Dynamism in Industrial Clusters in Developing Countries." *World Development* 27 (9): 1715–1734. doi:10.1016/S0305-750X(99)00073-X.

Brandt, Loren, Thomas G. Rawski, and John Sutton. 2008. "China's Industrial Development." In *China's Great Economic Transformation: Economics Development and Growth*, edited by Loren Brandt and Thomas G. Rawski, 1–26. Cambridge UK: Cambridge University Press.

Brandt, Loren, and Eric Thun. 2010. "The Fight for the Middle: Upgrading, Competition, and Industrial Development in China." *World Development* 38 (11): 1555–1574. doi:10.1016/j.worlddev.2010.05.003.

Brandt, Loren, Trevor Tombe, and Xiaodong Zhu. 2013. "Factor Market Distortions across Time, Space and Sectors in China." *Review of Economic Dynamics* 16 (1): 39–58. doi:10.1016/j.red.2012.10.002.

Callaert, J., M. du Plessis, J. Growels, C. Lecocq, T. Magerman, B. Peeters, X. Song, B. Van Looy, and C. Vereyen. 2011. *Patent Statistics at Eurostat: Methods for Regionalisation, Sector Allocation and Name Harmonisation*. *Eurostat*. Luxembourg.

Chen, Chung, Lawrence Chang, and Yimin Zhang. 1995. "The Role of Foreign Direct Investment in China's Post-1978 Economic Development." *World Development* 23 (4): 691–703. doi:10.1016/0305-750X(94)00143-M.

Chen, Kaihua, and Jiancheng Guan. 2011. "Mapping the Functionality of China's Regional Innovation Systems: A Structural Approach." *China Economic Review* 22 (1): 11–27. doi:10.1016/j.chieco.2010.08.002.

Chen, Kaihua, and Jiancheng Guan. 2012. “Measuring the Efficiency of China’s Regional Innovation Systems: Application of Network Data Envelopment Analysis (DEA).” *Regional Studies* 46 (3). Routledge: 355–377. doi:10.1080/00343404.2010.497479.

Chen, Kun, and Martin Kenney. 2007. “Universities/Research Institutes and Regional Innovation Systems: The Cases of Beijing and Shenzhen.” *World Development* 35 (6): 1056–1074. doi:10.1016/j.worlddev.2006.05.013.

Chen, Shiyi, Gary H. Jefferson, and Jun Zhang. 2011. “Structural Change, Productivity Growth and Industrial Transformation in China.” *China Economic Review* 22 (1): 133–150. doi:10.1016/j.chieco.2010.10.003.

Chen, Yun-Chung. 2007. “The Upgrading of Multinational Regional Innovation Networks in China.” *Asia Pacific Business Review* 13 (3). Routledge: 373–403. doi:10.1080/13602380701291941.

Cooke, Philip. 2001. “Regional Innovation Systems, Clusters, and the Knowledge Economy.” *Industrial and Corporate Change* 10 (4): 945–974. doi:10.1093/icc/10.4.945.

Cooke, Philip, Mikel Gomez Uranga, and Goio Etxebarria. 1997. “Regional Innovation Systems: Institutional and Organisational Dimensions.” *Research Policy* 26 (4-5): 475–491. doi:10.1016/S0048-7333(97)00025-5.

Dahlman, Carl J., and Jean-Eric Aubert. 2001. *China and the Knowledge Economy: Seizing the 21th Century*. Washington DC: The World Bank. doi:10.1596/0-8213-5005-6.

Deng, Ping. 2009. “Why Do Chinese Firms Tend to Acquire Strategic Assets in International Expansion?” *Journal of World Business* 44 (1): 74–84. doi:10.1016/j.jwb.2008.03.014.

Dernis, Hélène, and Dominique Guellec. 2001. “Using Patent Counts for Cross-Country Comparisons of Technology Output.” *STI Review* 27 (Special Issue on New Science and Technology Indicators): 129–146.

Dernis, Hélène, and Mosahid Khan. 2004. *Triadic Patent Families Methodology*. 2004/02. OECD Science, Technology and Industry Working Papers. Paris. doi:10.1787/443844125004.

Fagerberg, Jan, and Martin Srholec. 2008. “National Innovation Systems, Capabilities and Economic Development.” *Research Policy* 37 (9): 1417–1435. doi:http://dx.doi.org/10.5202/rei.v1i2.15.

Fan, Peilei. 2014. “Innovation in China.” *Journal of Economic Surveys* 28 (4): 725–745. doi:10.1111/joes.12083.

Felipe, J., U. Kumar, N. Usui, and A. Abdon. 2013. “Why Has China Succeeded? And Why It Will Continue to Do so.” *Cambridge Journal of Economics* 37 (4): 791–818. doi:10.1093/cje/bes058.

Frattini, Federico, and Giorgio Prodi. 2013a. “Industria E Crescita in Cina: Quali Evidenze, Quale Modello E Quali Criticità?” *L’industria* 2103 (3): 491–510. doi:10.1430/74605.

Frattini, Federico, and Giorgio Prodi. 2013b. “Industrial Clusters in China: Policy Tools for Further and More Balanced Development.” *European Review of Industrial Economics and Policy*, no. 5.

Fu, Xiaolan. 2008. “Foreign Direct Investment, Absorptive Capacity and Regional Innovation Capabilities: Evidence from China.” *Oxford Development Studies* 36 (1). Routledge: 89–110. doi:10.1080/13600810701848193.

Fu, Xiaolan, Carlo Pietrobelli, and Luc Soete. 2011. “The Role of Foreign Technology and Indigenous Innovation in the Emerging Economies: Technological Change and Catching-Up.” *World Development* 39 (7): 1204–1212. doi:10.1016/j.worlddev.2010.05.009.

Gabriele, Alberto. 2010. “The Role of the State in China’s Industrial Development: A Reassessment.” *Comparative Economic Studies* 52 (3): 325–350. doi:10.1057/ces.2010.11.

Girma, S., Y. Gong, and H. Gorg. 2008. “Foreign Direct Investment, Access to Finance, and Innovation Activity in Chinese Enterprises.” *The World Bank Economic Review* 22 (2): 367–382. doi:10.1093/wber/lhn009.

Griliches, Zvi. 1990. “Patent Statistics as Economic Indicators: A Survey.” *Journal of Economic Literature* 28 (4): 1661–1707.

Gu, Shulin, and Bengt-Åke Lundvall. 2006. “China’s Innovation System and the Move Toward Harmonious Growth and Endogenous Innovation.” *Innovation: Management, Policy & Practice* 8 (Innovation in China: Harmonious Transformation?): 1–26.

Guan, Jiancheng, and Shunzhong Liu. 2005. “Comparing Regional Innovative Capacities of PR China Based on Data Analysis of the National Patents.” *International Journal of Technology Management* 32 (3-4): 225–245.

Hair, Joseph F. Jr., William C. Black, Barry J. Babin, and Rolph E. Anderson. 2009. *Multivariate Data Analysis*. Seventh. Upper Saddle River NJ: Prentice Hall.

He, Canfei, Yehua Dennis Wei, and Xiuzhen Xie. 2008. “Globalization, Institutional Change, and Industrial Location: Economic Transition and Industrial Concentration in China.” *Regional Studies*. doi:10.1080/00343400701543272.

He, Dong, and Wenlang Zhang. 2010. “How Dependent Is the Chinese Economy on Exports and in What Sense Has Its Growth Been Export-Led?” *Journal of Asian Economics* 21 (1): 87–104. doi:10.1016/j.asieco.2009.04.005.

Heilmann, Sebastian. 2008. “Policy Experimentation in China’s Economic Rise.” *Studies in Comparative International Development* 43 (1): 1–26. doi:10.1007/s12116-007-9014-4.

Hemphill, Thomas A., and George O. White. 2013. “China’s National Champions: The Evolution of a National Industrial Policy-Or a New Era of Economic Protectionism?” *Thunderbird International Business Review* 55 (2): 193–212. doi:10.1002/tie.21535.

Hu, Albert Guangzhou. 2007. “Technology Parks and Regional Economic Growth in China.” *Research Policy* 36 (1): 76–87. doi:10.1016/j.respol.2006.08.003.

Keller, Wolfgang. 2004. “International Technology Diffusion.” *Journal of Economic Literature* 42 (3): 752–782. doi:10.1257/0022051042177685.

Kroll, Henning, and Ingo Liefner. 2008. “Spin-off Enterprises as a Means of Technology Commercialisation in a Transforming economy—Evidence from Three Universities in China.” *Technovation* 28 (5): 298–313. doi:10.1016/j.technovation.2007.05.002.

Lall, Sanjaya. 1992. “Technological Capabilities and Industrialization.” *World Development* 20 (2): 165–186. doi:10.1016/0305-750X(92)90097-F.

Lau, Antonio K W, and William Lo. 2015. “Regional Innovation System, Absorptive Capacity and Innovation Performance: An Empirical Study.” *Technological Forecasting & Social Change* 92 (March): 99–114. doi:10.1016/j.techfore.2014.11.005.

Lau, Lawrence J., Yingyi Qian, and Gerard Roland. 2000. “Reform Without Losers: An Interpretation of China’s Dual-Track Approach to Transition.” *Journal of Political Economy* 108 (1). The University of Chicago Press: 120. doi:10.1086/262113.

Li, Weiye, and Louis Putterman. 2008. “Reforming China’s SOEs: An Overview.” *Comparative Economic Studies* 50 (3). Nature Publishing Group: 353–380. doi:10.1057/ces.2008.31.

Li, Xibao. 2009. “China’s Regional Innovation Capacity in Transition: An Empirical Approach.” *Research Policy* 38 (2): 338–357. doi:10.1016/j.respol.2008.12.002.

Li, Xibao. 2012. “Behind the Recent Surge of Chinese Patenting: An Institutional View.” *Research Policy* 41 (1): 236–249. doi:10.1016/j.respol.2011.07.003.

Lin, Justin Yifu, and Yan Wang. 2012. “China’s Integration with the World: Development as a Process of Learning and Industrial Upgrading.” *China Economic Policy Review* 1 (1): 1–33. doi:10.1142/S179396901250001X.

Lundvall, Bengt-Åke, K.J. Joseph, Cristina Chaminade, and Jan Vang, eds. 2009. *Handbook of Innovation Systems and Developing Countries: Building Domestic Capabilities in a Global Setting*. Hardcover. Cheltenham UK: Edward Elgar Publishing.

Manly, Bryan F.J. 2004. *Multivariate Statistical Methods: A Primer*. Third. Boca Raton FL: Chapman and Hall/CRC.

Naughton, Barry. 2007. *The Chinese Economy: Transitions and Growth*. Cambridge MA: MIT Press.

OECD. 2009. *OECD Patent Statistics Manual*. Paris: OECD Publishing.

Rawski, Thomas G. 1994. “Chinese Industrial Reform: Accomplishments, Prospects, and Implications.” *The American Economic Review* 84 (2). American Economic Association: 271–275.

Rizzo, Ugo, Francesco Nicolli, and Laura Ramaciotti. 2013. “The Heterogeneity of the Development Process of New Technology-Based Firms. Implication for Innovation Policies.” *Journal of the Knowledge Economy* 5 (1): 114–132. doi:10.1007/s13132-013-0170-7.

Rodrik, Dani. 2006. “What’s So Special about China's Exports?” *China & World Economy* 14 (5): 1–19. doi:10.1111/j.1749-124X.2006.00038.x.

Srholec, Martin. 2011. “A Multilevel Analysis of Innovation in Developing Countries.” *Industrial and Corporate Change* 20 (6): 1539–1569. doi:10.1093/icc/dtr024.

Sun, Sizhong. 2012. “The Role of FDI in Domestic Exporting: Evidence from China.” *Journal of Asian Economics* 23 (4): 434–441. doi:10.1016/j.asieco.2012.03.004.

Sun, Yifei. 2002. “Sources of Innovation in China’s Manufacturing Sector: Imported or Developed in-House?” *Environment and Planning A* 34 (6): 1059–1072.

Sun, Yutao, and Fengchao Liu. 2010. “A Regional Perspective on the Structural Transformation of China’s National Innovation System since 1999.” *Technological Forecasting and Social Change* 77 (8): 1311–1321. doi:10.1016/j.techfore.2010.04.012.

Tian, Qunjian. 2004. “China Develops Its West: Motivation, Strategy and Prospect.” *Journal of Contemporary China* 13 (41): 611–636. doi:10.1080/1067056042000281404.

Wang, Cassandra C., and George C. S. Lin. 2012. “Dynamics of Innovation in a Globalizing China: Regional Environment, Inter-Firm Relations and Firm Attributes.” *Journal of Economic Geography* 13 (3): 397–418. doi:10.1093/jeg/lbs019.

World Bank. 2010. *Innovation Policy: A Guide for Developing Countries*. Washington DC: The World Bank. doi:10.1596/978-0-8213-8269-1.

Wu, Weiping. 2007a. “Cultivating Research Universities and Industrial Linkages in China: The Case of Shanghai.” *World Development* 35 (6): 1075–1093. doi:10.1016/j.worlddev.2006.05.011.

Wu, Weiping. 2007b. “State Policies, Enterprise Dynamism, and Innovation System in Shanghai, China.” *Growth and Change* 38 (4): 544–566. doi:10.1111/j.1468-2257.2007.00387.x.

Zeng, Douglas Zhihua. 2010. “How Do Special Economic Zones and Industrial Clusters Drive China’s Rapid Growth?” In *Building Engines for Growth and Competitiveness in China: Experience with Special Economic Zones and Industrial Clusters*, edited by Douglas Zhihua Zeng, 1–54. Washington DC: The World Bank.

Zhang, Xiaobo, and Kong Yam Tan. 2007. “Incremental Reform and Distortions in China’s Product and Factor Markets.” *World Bank Economic Review* 21 (2): 279–299. doi:10.1093/wber/lhm002.

Zhao, S.L., L. Cacciolatti, S.H. Lee, and W. Song. 2015. “Regional Collaborations and Indigenous Innovation Capabilities in China: A Multivariate Method for the Analysis of Regional Innovation Systems.” *Technological Forecasting and Social Change* 94 (October): 202–220. doi:10.1016/j.techfore.2014.09.014.

Appendix 1. Patent filings at the EPO and SIPO compared

Data at the SIPO from the China Data On Line, Yearbooks Database are available at a provincial level and after 1985 only. Therefore, Tables 4(a) and 4(b) refer to this shorter period and broader regional detail. Nonetheless, with few exceptions, and despite a relatively few patents at the EPO, correlations between the number of documents filed at the EPO and SIPO from Chinese applicants tend to be very high both in time and regions. This result suggests that an empirical strategy focusing on the EPO applications is statistically, and not only conceptually, robust.

[Tables 4(a) and 4(b) about here]

Appendix 2. Robustness

As a control, this Appendix presents an analysis that is alternative to the main clustering procedure in Section 4. Differences rely on the treatment of missing values. Here, missing values for all indexes are replaced with zeros when prefectures recorded at least one patent application at the EPO between 1981 and 2009. Formerly, in Section 4 the replacement rule was at least one patent application at the EPO within a reform stage, not included otherwise. Thus, the analysis now includes much more zeros, and indexes average value at the country level is of course lower, especially in Stages 1 and 2 (Figure 7). Despite of this change, the expectation is that results obtained here are very similar to those in Section 4 but comprising also a wide group of prefectures with very poor values. However, the size of this additional group is expected to decrease over time, and the overall number of prefectures is now the same across reform stages (200). For everything else, the procedure follows the same steps as in the main analysis.

[Figure 7 about here]

At the first (hierarchical) step, results suggest a clearly identifiable eight-cluster solution for the first stage and a five-cluster solution for the third (agglomeration coefficients change from 9.4 to 20.6 and from 5.7 to 8.4 respectively). Conversely, the choice between a six- and seven-cluster solution for the second period is unclear, but the first is preferred because providing groups with an easier interpretation. Again, the analysis proceeds taking the centroids obtained at the first step as the initial seeds in a (second) K-means non-hierarchical clustering procedure that generates the results shown in Tables 5(a) to 5(c), one for each reform stage.

[Table 5(a), 5(b) and 5(c) about here]

As expected, a new group is emerging, notably in the period 1981-1992 (7). This group exhibits poor values and consists of prefectures with very low or null patent counts. Moreover, at least three other relevant results confirm those from the main analysis. First, centroid values generally increase stage-by-stage coherently to the distributional pattern previously shown in Figures 6(a) to 6(c), due to a patenting propensity that grows over time, including among the lower performing prefectures. Second, the number of groups reduces over time and suggests that innovation activities went generally reinforcing reinforcement. Final, the number of prefectures with poor activities is still decreasing over time. Therefore, this alternative analysis confirms the main evidences presented in Section 4. Unfortunately, it weakens in part the possibility of capturing the diffusion of innovation activities because several prefectures that exhibit low index values (0, Table 6) are now unable to join the group where *INV* prevails (1).

[Table 6 about here]

Table 1(a). Cluster descriptions: first reform stage (1981-1992).

Clusters		1	2	3	4	5	6	7	Total
	APP	0.04388	0.00844	0.10012	0.04687	0.04674	0.00000	0.00142	
Centroids	*BOTH*	0.28304	0.02849	0.40073	0.26693	0.19265	0.04609	0.12643	
	INV	0.01923	0.04333	0.42221	0.22465	0.45290	0.16929	0.06444	
Cases		2	23	1	2	1	5	6	40

Table 1(b). Cluster descriptions: second reform stage (1993-2001).

Clusters		1	2	3	4	5	6	Total
	APP	0.07284	0.04466	0.22711	0.04130	0.02932	0.03134	
Centroids	*BOTH*	0.33583	0.23242	0.52853	0.15456	0.04047	0.07535	
	INV	0.41671	0.10675	0.24434	0.58191	0.06797	0.25626	
Cases		7	11	1	3	50	15	87

Table 1(c). Cluster descriptions: third reform stage (2002-2009).

Clusters		1	2	3	4	5	Total
	APP	0.11321	0.40605	0.03355	0.03259	0.10201	
Centroids	*BOTH*	0.65966	0.26426	0.19315	0.05696	0.40683	
	INV	0.16462	0.07968	0.53891	0.10671	0.24115	
Cases		14	5	16	114	38	187

Table 2. Table of results converted into adjusted clusters (ACL).

Adjusted clusters (*ACL*)		Cluster results			Number of cases		
Code	Description	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
0	No activity (not included)				305	258	158
1	Prevalence of *INV*	2, 5, 6	4, 5, 6	3, 4	29	68	130
2	Prevalence of *INV* and *BOTH*	3, 4	1	5	3	7	38
3	Prevalence of *BOTH*	1, 7	2, 3	1	8	12	14
4	Prevalence of *BOTH* and *APP*			2	0	0	5
5	Prevalence of *APP*				0	0	0

Table 3. Frequency distribution of the adjusted clusters (ACL) by reform stage in selected groups of Chinese prefectures (percentage share).

		ACL	Stage 1	Stage 2	Stage 3
SEZs	Hainan, Shanghai, Shantou, Shenzhen, Tianjin, Xiamen and Zhuhai	0	28.6	14.3	14.3
		1	42.9	28.6	28.6
		2	14.3	28.6	28.6
		3	14.3	28.6	28.6
		4	0.0	0.0	0.0
		5	0.0	0.0	0.0
Provinces	Fujian (excluding Xiamen)	0	75.0	75.0	12.5
		1	25.0	25.0	75.0
		2	0.0	0.0	12.5
		3	0.0	0.0	0.0
		4	0.0	0.0	0.0
		5	0.0	0.0	0.0
	Guangdong (excluding Shantou, Shenzhen and Zhuhai)	0	88.9	61.1	27.8
		1	5.6	33.3	38.9
		2	5.6	0.0	33.3
		3	0.0	5.6	0.0
		4	0.0	0.0	0.0
		5	0.0	0.0	0.0
	Jiangsu	0	92.3	53.8	0.0
		1	7.7	23.1	38.5
		2	0.0	7.7	30.8
		3	0.0	15.4	15.4
		4	0.0	0.0	15.4
		5	0.0	0.0	0.0
	Zhejiang	0	72.7	54.5	0.0
		1	18.2	27.3	45.5
		2	0.0	9.1	18.2
		3	9.1	9.1	27.3
		4	0.0	0.0	9.1
		5	0.0	0.0	0.0

Table 4(a). Correlation between patent applications to the EPO and SIPO by province (variability over the years 1985-2009). Chinese applicants. Source: authors' arrangement from the China Data On Line, Yearbooks Database and OECD statistics.

Region	SIPO documents	EPO documents	EPO / 1,000 SIPO	Correlation
China	4,828,786	23,563	4.88	0.99
Guangdong	757,272	6,847	9.04	0.95
Jiangsu	603,856	473	0.78	0.95
Zhejiang	499,642	411	0.82	0.99
Shandong	367,091	211	0.58	0.97
Shanghai	348,893	680	1.95	0.97
Beijing	316,762	1,317	4.16	0.98
Liaoning	204,649	120	0.59	0.92
Sichuan	165,665	107	0.64	0.86
Hubei	142,096	88	0.62	0.93
Henan	124,651	31	0.25	0.72
Tianjin	123,289	109	0.89	0.87
Hunan	121,310	101	0.83	0.83
Fujian	112,683	135	1.20	0.90
Hebei	96,541	53	0.55	0.82
Heilongjiang	84,719	17	0.20	0.40
Shaanxi	79,060	100	1.26	0.95
Anhui	64,791	46	0.70	0.86
Chongqing	61,462	59	0.97	0.85
Jilin	59,861	41	0.68	0.50
Jiangxi	40,785	41	1.01	0.83
Guangxi	40,304	11	0.26	0.47
Shanxi	39,386	19	0.47	0.60
Yunnan	37,036	11	0.30	0.73
Guizhou	26,755	18	0.68	0.37
Xinjiang	24,868	11	0.42	0.73
Inner Mongolia	23,746	17	0.72	0.56
Gansu	19,843	14	0.68	0.25
Hainan	8,603	14	1.66	0.42
Ningxia	8,482	10	1.18	0.81
Qinghai	4,053	1	0.25	0.38
Tibet	1,086	5	4.60	-0.23

Table 4(b). Correlation between patent applications to the EPO and SIPO by year (variability over provinces). Chinese applicants (1985-2009). Source: authors' arrangement from the China Data On Line, Yearbooks Database and OECD statistics.

Year	SIPO documents	EPO documents	EPO / 1,000 SIPO	Correlation
1985	9,411	108	11.44	0.80
1986	8,945	83	9.27	0.51
1987	14,315	105	7.30	0.49
1988	7,328	114	15.58	0.51
1989	27,367	130	4.74	0.71
1990	36,585	154	4.20	0.52
1991	45,395	176	3.88	0.63
1992	61,788	190	3.07	0.58
1993	68,153	168	2.46	0.62
1994	67,807	172	2.54	0.61
1995	68,880	171	2.49	0.48
1996	39,725	207	5.20	0.52
1997	90,071	259	2.88	0.50
1998	96,233	350	3.63	0.44
1999	109,958	468	4.26	0.54
2000	140,339	555	3.95	0.78
2001	165,773	779	4.70	0.87
2002	205,544	1,022	4.97	0.86
2003	251,238	1,222	4.86	0.87
2004	278,943	1,503	5.39	0.85
2005	383,157	2,178	5.68	0.79
2006	470,342	2,442	5.19	0.78
2007	586,734	3,209	5.47	0.68
2008	717,144	3,270	4.56	0.56
2009	877,611	4,531	5.16	0.52

Table 5(a). Cluster descriptions: first reform stage (1981-1992).

Clusters		1	2	3	4	5	6	7	8	Total
Centroids	*APP*	0.04388	0.00488	0.10012	0.04674	0.00000	0.09374	0.00079	0.00000	
	BOTH	0.28304	0.09834	0.40073	0.19265	0.04609	0.25182	0.00077	0.28205	
	INV	0.01923	0.03287	0.42221	0.45290	0.16929	0.19289	0.00543	0.25641	
Cases		2	13	1	1	5	1	176	1	200

Table 5(b). Cluster descriptions: second reform stage (1993-2001).

Clusters		1	2	3	4	5	6	Total
	APP	0.00795	0.13944	0.04991	0.01978	0.09189	0.00569	
Centroids	*BOTH*	0.00782	0.15167	0.22896	0.18924	0.39110	0.06458	
	INV	0.01647	0.21682	0.56556	0.08726	0.36145	0.25450	
Cases		155	7	5	15	5	13	200

Table 5(c). Cluster descriptions: third reform stage (2002-2009).

Clusters		1	2	3	4	5	Total
	APP	0.11321	0.40605	0.03232	0.02925	0.10073	
Centroids	*BOTH*	0.16462	0.07968	0.54143	0.09578	0.24781	
	INV	0.65966	0.26426	0.17624	0.05113	0.40786	
Cases		14	5	15	127	39	200

Table 6. Table of results converted into adjusted clusters.

Adjusted clusters (*ACL*)		Cluster results			Number of cases		
Code	Description	Stage 1	Stage 2	Stage 3	Stage 1	Stage 2	Stage 3
N	No activity (not included)				145	145	145
0	Poor activity	2, 7	1	4	189	155	127
1	Prevalence of *INV*	4, 5	2, 3, 6	1	6	25	14
2	Prevalence of *INV* and *BOTH*	3, 6, 8	5	5	3	5	39
3	Prevalence of *BOTH*	1	4	3	2	15	15
4	Prevalence of *BOTH* and *APP*			2	0	0	5
5	Prevalence of *APP*				0	0	0

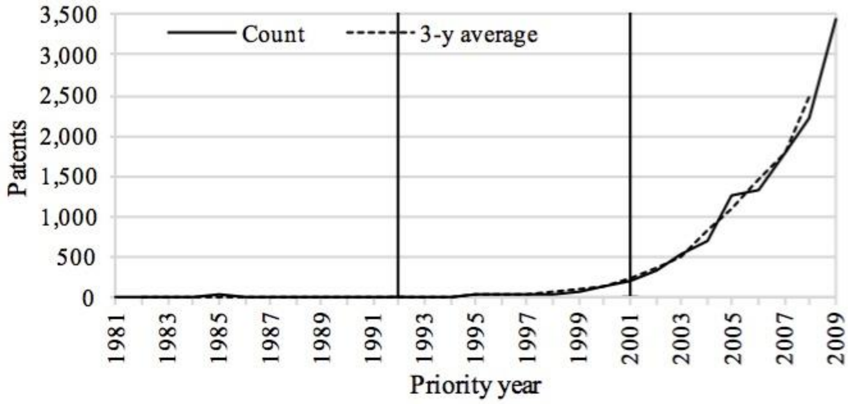

Figure 1. Patent applications to the European Patent Office by year (1981-2009). Chinese inventors. Source: authors' arrangement from the OECD REGPAT database, January 2014.

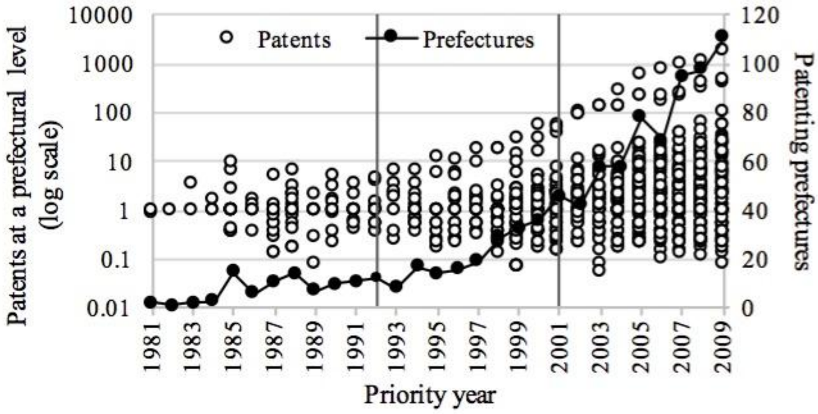

Figure 2. Patent distribution across Chinese prefectures by year (1981-2009). Patent applications to the European Patent Office (logarithmic scale). Inventors from China. Source: authors' arrangement from the OECD REGPAT database, January 2014.

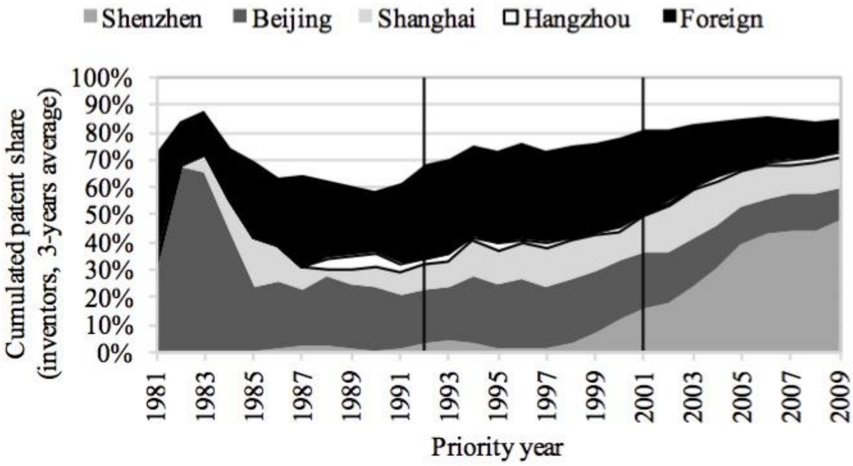

Figure 3(a). Cumulated patent shares by year (1981-2009). Patent applications to the European Patent Office from China. Inventors from the four prefectures with highest patent counts and abroad. Source: authors' arrangement from the OECD REGPAT database, January 2014.

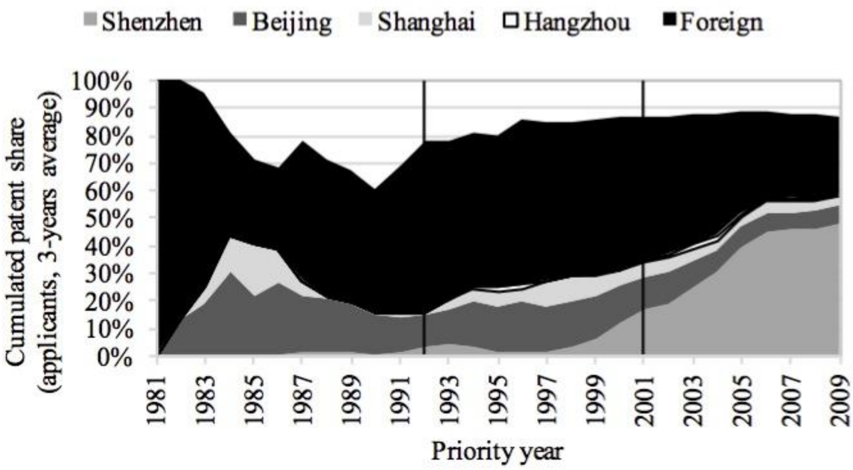

Figure 3(b). Cumulated patent shares by year (1981-2009). Patent applications to the European Patent Office from China. Applicants from the four prefectures with highest patent counts and abroad. Source: authors' arrangement from the OECD REGPAT database, January 2014.

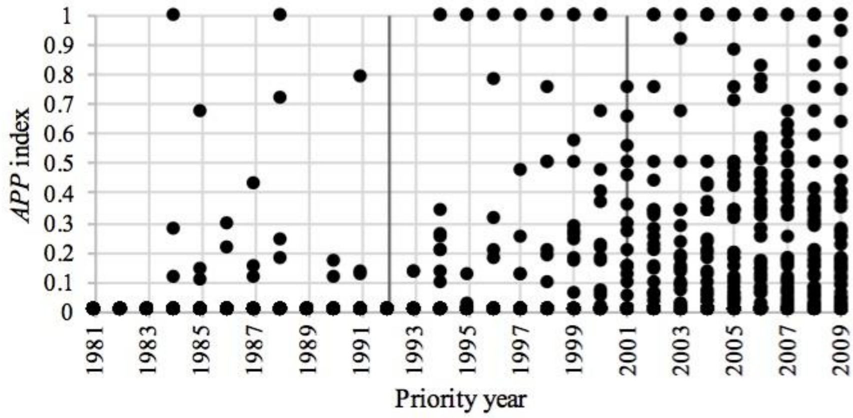

Figure 4(a). APP index distribution across Chinese prefectures by year (1981-2009). Source: authors' arrangement from the OECD REGPAT database, January 2014.

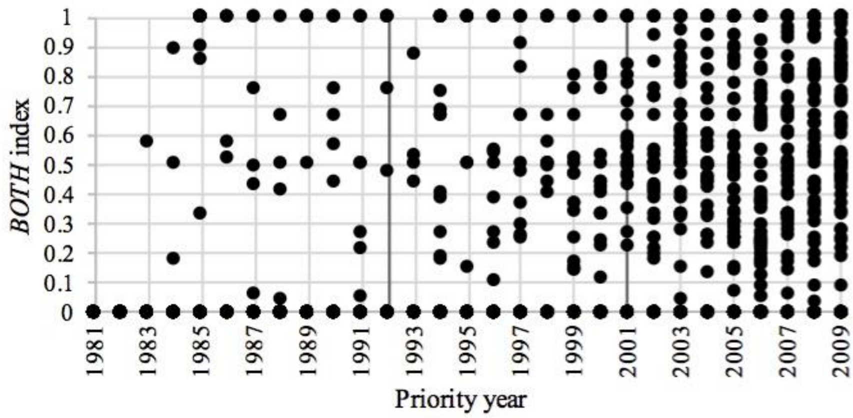

Figure 4(b). BOTH index distribution across Chinese prefectures by year (1981-2009). Source: authors' arrangement from the OECD REGPAT database, January 2014.

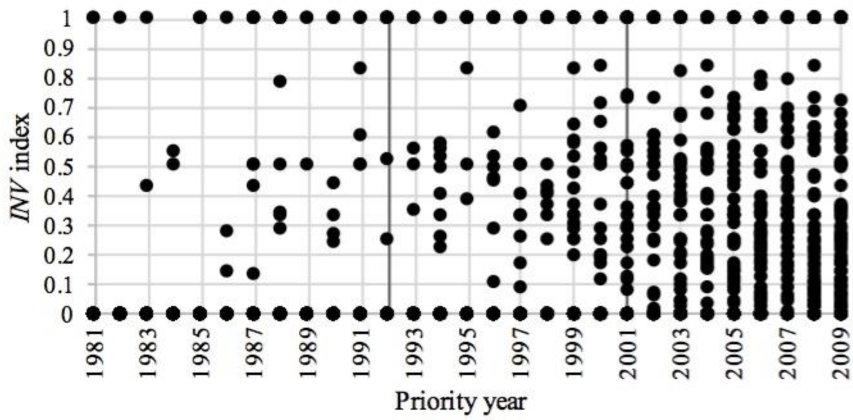

Figure 4(c). INV index distribution over Chinese prefectures by year (1981-2009). Source: authors' arrangement from the OECD REGPAT database, January 2014.

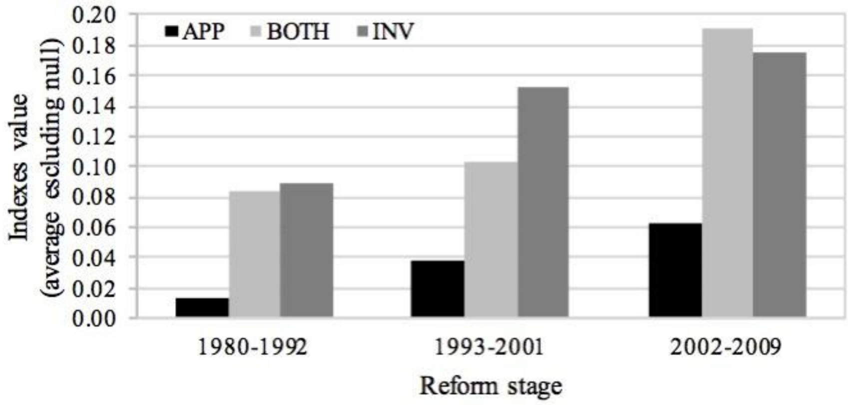

Figure 5. Indexes: country average value by reform stage (excluding null values). Source: authors' arrangement from the OECD REGPAT database, January 2014.

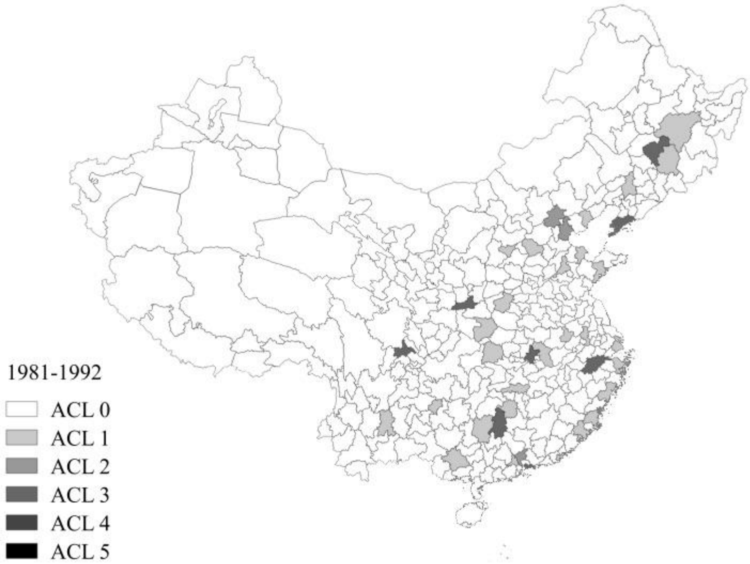

Figure 6(a). Adjusted clusters (ACL) during the first reform stage (1981-1992).

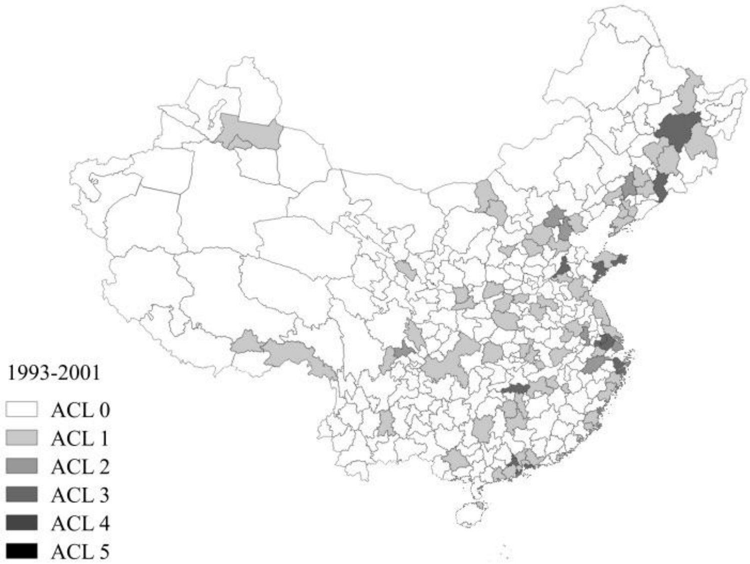

Figure 6(b). Adjusted clusters (ACL) during the second reform stage (1993-2001).

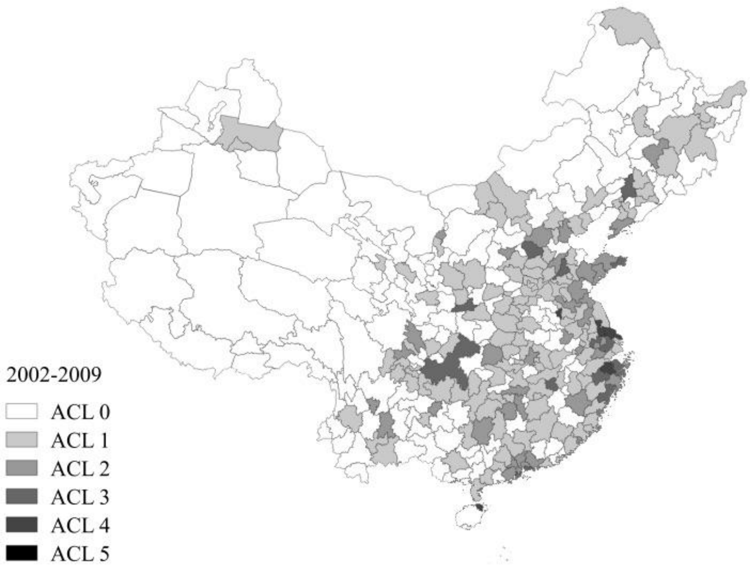

Figure 6(c). Adjusted clusters (ACL) during the third reform stage (2002-2009).

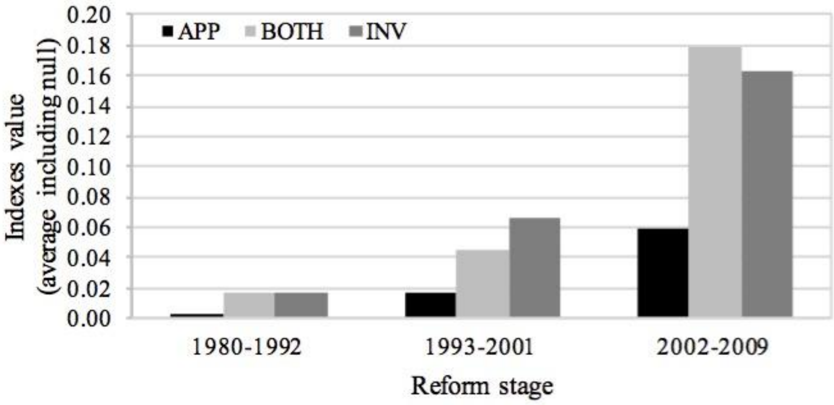

Figure 7. Indexes: country average value by reform stage (including null values as zeros). Source: authors' arrangement from the OECD REGPAT database, January 2014.